View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Computers and Mathematics with Applications 62 (2011) 4605–4613

Contents lists available at SciVerse ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

On the global Krylov subspace methods for solving general coupled matrix equations

Fatemeh Panjeh Ali Beik [a], Davod Khojasteh Salkuyeh [b,*]

[a] Department of Mathematics, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran

[b] Department of Mathematics, University of Mohaghegh Ardabili, P.O. Box 179, Ardabil, Iran

ARTICLE INFO

Article history:
Received 11 August 2011
Received in revised form 17 October 2011
Accepted 17 October 2011

Keywords:
Linear matrix equation
Krylov subspace
Global FOM
Global GMRES

ABSTRACT

In the present paper, we propose the global full orthogonalization method (Gl-FOM) and global generalized minimum residual (Gl-GMRES) method for solving large and sparse general coupled matrix equations

$$\sum_{j=1}^{p} A_{ij}X_jB_{ij} = C_i, \quad i = 1, \dots, p,$$

where $A_{ij} \in \mathbb{R}^{m\times m}$, $B_{ij} \in \mathbb{R}^{n\times n}$, $C_i \in \mathbb{R}^{m\times n}$, $i, j = 1, 2, \dots, p$, are given matrices and $X_i \in \mathbb{R}^{m\times n}$, $i = 1, 2, \dots, p$, are the unknown matrices. To do so, first, a new inner product and its corresponding matrix norm are defined. Then, using a linear operator equation and new matrix product, we demonstrate how to employ Gl-FOM and Gl-GMRES algorithms for solving general coupled matrix equations. Finally, some numerical experiments are given to illustrate the validity and applicability of the results obtained in this work.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

In this paper, we consider the general coupled matrix equations of the form

$$\sum_{j=1}^{p} A_{ij}X_jB_{ij} = C_i, \quad i = 1, \dots, p, \tag{1.1}$$

where $A_{ij} \in \mathbb{R}^{m\times m}$, $B_{ij} \in \mathbb{R}^{n\times n}$, and $C_i \in \mathbb{R}^{m\times n}$, $i, j = 1, 2, \dots, p$, are large and sparse matrices, $X_i \in \mathbb{R}^{m\times n}$, $i = 1, 2, \dots, p$, are the unknown matrices. Such problems arise in linear control and filtering theory for continues or discrete-time large-scale dynamical systems. They also play an important role in image restoration and other problems; for more details see [1–5] and the references therein.

Many investigated matrix equations in the literature can be considered as special cases of (1.1). For example, Bouhamidi and Jbilou [1] have considered the generalized Sylvester matrix equation

$$\sum_{j=1}^{p} A_jXB_j = C, \tag{1.2}$$

* Corresponding author. Tel.: +98 451 5514702; fax: +98 451 5514701.
E-mail addresses: f.beik@vru.ac.ir (F.P.A. Beik), khojaste@uma.ac.ir, salkuyeh@yahoo.com (D.K. Salkuyeh).

0898-1221/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2011.10.043

and proposed a Krylov subspace method for solving (1.2). In [6], Li and Wang proposed an iterative algorithm for the minimal norm least squares solution to (1.2). Chang and Wang [7] have presented necessary and sufficient conditions for the existence and the expressions for the symmetric solutions of the matrix equations

$$\begin{cases} AX + YA = C, \\ AXA^T + BYB^T = C, \end{cases}$$

and

$$(A^T XA, B^T XB) = (C, D).$$

In [8], Wang et al. have given necessary and sufficient conditions for the existence of constant solutions with bi(skew)symmetric constrains to the matrix equations

$$A_i X - YB_i = C_i, \quad i = 1, 2, \ldots, s,$$

and

$$A_i XB_i - C_i YD_i = E_i, \quad i = 1, 2, \ldots, s.$$

A good survey of the methods to solve special cases of the general coupled matrix (1.1) can be found in [9].

It is easy to see that the general coupled matrix (1.1) is equivalent to

$$\sum_{j=1}^{p} (B_{ij}^T \otimes A_{ij}) \mathrm{vec}(X_j) = \mathrm{vec}(C_i), \quad i = 1, \ldots, p, \tag{1.3}$$

where $\otimes$ denotes the Kronecker product operator and $\mathrm{vec}(Z) = (z_1^T, z_2^T, \ldots, z_m^T)^T$ for $Z = (z_1, z_2, \ldots, z_n) \in \mathbb{R}^{m\times n}$. Obviously, the coefficient matrix of the linear system (1.3) is of order pmn and can be solved by iterative methods such as the methods based on the Krylov subspace methods like the GMRES [10]. Evidently, the size of the linear system (1.3) would be huge even for moderate values of m, n and p. Therefore, it is more preferable to employ an iterative method for solving the original system (1.1) instead of the linear system (1.3). Note that system (1.1) has a unique solution if and only if the coefficient matrix of the linear system (1.3) is nonsingular. Throughout this paper we assume that the system (1.1) has a unique solution.

In [9], Dehghan and Hajarian have presented an iterative method to solve the general coupled matrix equations (1.1) over generalized bisymmetric matrix group $(X_1, X_2, \ldots, X_p)$. In [11], a gradient based algorithm and a least square based iterative algorithm have been presented for solving (1.2). Ding and Chen [12] used the hierarchical identification principle to construct iterative solutions to the coupled linear matrix equation (1.1). In [13], Zhou et al. proposed an iterative method for finding weighted least squares solutions to system (1.1). A gradient based iterative algorithm for solving coupled matrix equations has been presented by Zhou et al. in [14]. Recently, Zhang in [4] has extended the CGNE [15] and Bi-CGSTAB [15] algorithms to solve (1.1).

In [2], the global Krylov subspace methods have been originally presented for solving a linear system of equations with multiple right-hand sides. It is well-known that the global Krylov subspace methods outperform other iterative methods for solving such systems when the coefficient matrix is large and nonsymmetric. On the other hand, the global Krylov subspace methods are also effective when applied for solving large and sparse linear matrix equations; for more details see [1,16,17] and the references therein. Therefore, we are interested in employing the global Krylov subspaces for solving (1.1) when the coefficient matrices are large and sparse. To do so, we first define the linear operator $\mathcal{M}$ as follows

$$\mathcal{M} : \mathbb{R}^{m\times n} \times \cdots \times \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times pn},$$
$$X = (X_1, X_2, \ldots, X_p) \to \mathcal{M}(X) = (\mathcal{A}_1(X), \mathcal{A}_2(X), \ldots, \mathcal{A}_p(X)),$$

where

$$\mathcal{A}_i(X) = \sum_{j=1}^{p} A_{ij} X_j B_{ij}, \quad i = 1, 2, \ldots, p.$$

Using the linear operator $\mathcal{M}$, we rewrite Eq. (1.1) as

$$\mathcal{M}(X) = C, \tag{1.4}$$

where $C = (C_1, C_2, \ldots, C_p)$. In the next sections, we utilize the linear matrix operator $\mathcal{M}$ to present Gl-FOM and Gl-GMRES algorithms for solving (1.1). More precisely, we focus on the solution of Eq. (1.4) instead of Eq. (1.1).

The rest of the paper is organized as follows. In Section 2, we first recall some necessary definitions and notations, then a new inner product is presented. We also introduce a new matrix product and give some of its properties. Section 3 is devoted to employing the Gl-FOM and Gl-GMRES algorithms for solving Eq. (1.4). In Section 4, some numerical experiments are given to show the efficiency of the proposed algorithms. Finally, the paper finishes with a brief conclusion in Section 5.

2. Preliminaries

In this section, we review some notations and definitions which are utilized throughout this paper. Moreover, we introduce some new concepts which are useful for presenting the Gl-FOM and Gl-GMRES algorithms for solving Eq. (1.4).

For two matrices Y and Z in $\mathbb{R}^{m\times n}$, the inner product $\langle Y, Z\rangle_F$ is defined as $\langle Y, Z\rangle_F = \mathrm{tr}(Y^T Z)$, the associate norm is the Frobenius norm denoted by $\|.\|_F$.

Definition 2.1 (*Bouyouli et al. [18]*)**.** Let $A = [A_1, A_2, \ldots, A_p]$ and $B = [B_1, B_2, \ldots, B_\ell]$ be matrices of dimensions $n \times ps$ and $n \times \ell s$, respectively, where A_i and B_j are $n \times s$ matrices. Then the matrix $A^T \diamond B = [(A^T \diamond B)_{ij}]_{p\times\ell}$ is defined by

$$(A^T \diamond B)_{ij} = \langle A_i, B_j\rangle_F.$$

In the following, we define a new inner product and its corresponding matrix norm which are used for deriving our further results in this paper.

Definition 2.2. Assume that $\overline{X} = (\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_p)$ and $\widetilde{X} = (\widetilde{X}_1, \widetilde{X}_2, \ldots, \widetilde{X}_p)$ are in $\mathbb{R}^{m\times pn}$. We define the inner product $\langle\cdot,\cdot\rangle$ as follows:

$$\langle \overline{X}, \widetilde{X}\rangle = \mathrm{tr}(\overline{X}^T \diamond \widetilde{X}). \tag{2.1}$$

Remark 2.3. For $X = (\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_p)$ in $\mathbb{R}^{m\times pn}$, the norm of X is defined by $\|X\|^2 = \mathrm{tr}(X^T \diamond X)$. Throughout this paper, a set of matrices in $\mathbb{R}^{m\times pn}$ is said to be orthonormal if it is orthonormal with respect to the scalar product (2.1).

Now, we introduce a new product denoted by $\circledcirc$ and defined as follows:

Definition 2.4. Let $A = [A^{(1)}, A^{(2)}, \ldots, A^{(k)}]$, $B = [B^{(1)}, B^{(2)}, \ldots, B^{(\ell)}]$ be $m \times kpn$ and $m \times \ell pn$ matrices, respectively, where $A^{(i)} = [A_1^{(i)}, A_2^{(i)}, \ldots, A_p^{(i)}]$, $B^{(s)} = [B_1^{(s)}, B_2^{(s)}, \ldots, B_p^{(s)}]$ and $A_j^{(i)}, B_j^{(s)} \in \mathbb{R}^{m\times n}$ for $i = 1, 2, \ldots, k$, $s = 1, 2, \ldots, \ell$ and $j = 1, 2, \ldots, p$. The $k \times \ell$ matrix $A^T \circledcirc B$ is defined by:

$$A^T \circledcirc B = \begin{pmatrix} \mathrm{tr}((A^{(1)})^T \diamond B^{(1)}) & \mathrm{tr}((A^{(1)})^T \diamond B^{(2)}) & \ldots & \mathrm{tr}((A^{(1)})^T \diamond B^{(\ell)}) \\ \mathrm{tr}((A^{(2)})^T \diamond B^{(1)}) & \mathrm{tr}((A^{(2)})^T \diamond B^{(2)}) & \ldots & \mathrm{tr}((A^{(2)})^T \diamond B^{(\ell)}) \\ \vdots & \vdots & \vdots & \vdots \\ \mathrm{tr}((A^{(k)})^T \diamond B^{(1)}) & \mathrm{tr}((A^{(k)})^T \diamond B^{(2)}) & \ldots & \mathrm{tr}((A^{(k)})^T \diamond B^{(\ell)}) \end{pmatrix}.$$

It is not difficult to establish the following remarks.

Remarks. (i) If $X = (X_1, X_2, \ldots, X_p) \in \mathbb{R}^{m\times pn}$, then $X^T \circledcirc X = \|X\|^2$.

(ii) The matrix $A = (A^{(1)}, A^{(2)}, \ldots, A^{(k)})$ is called orthonormal if and only if

$$A^T \circledcirc A = I_k.$$

(iii) Let the matrices A, B be defined as before and $L \in \mathbb{R}^{\ell\times k}$. Then

$$A^T \circledcirc (B((L \otimes I_p) \otimes I_n)) = (A^T \circledcirc B)L. \tag{2.2}$$

(iv) Let $A, B, C \in \mathbb{R}^{m\times kpn}$, then
   (a) $(A + B)^T \circledcirc C = A^T \circledcirc C + B^T \circledcirc C$.
   (b) $A^T \circledcirc (B + C) = A^T \circledcirc B + A^T \circledcirc C$.
   (c) $(A^T \circledcirc B)^T = B^T \circledcirc A$.

3. Implementing global Krylov subspace methods

In this section, we utilize Gl-FOM and Gl-GMRES algorithms to solve Eq. (1.4) which is equivalent to Eq. (1.1).

Suppose that $X^{(0)} = (X_1^{(0)}, X_2^{(0)}, \ldots, X_p^{(0)})$ in $\mathbb{R}^{m\times pn}$ is a given initial approximate solution and consider the Eq. (1.4). As a natural way, we define the matrix Krylov subspace as follows

$$\mathcal{K}_k(\mathcal{M}, R^{(0)}) = \mathrm{span}\left\{R^{(0)}, \mathcal{M}(R^{(0)}), \ldots, \mathcal{M}^{k-1}(R^{(0)})\right\}, \tag{3.1}$$

where $R^{(0)} = C - \mathcal{M}(X^{(0)})$.

3.1. Global Arnoldi process

In this subsection, we employ the global Arnoldi process to construct an orthonormal basis for the matrix Krylov subspace defined by (3.1).

Algorithm 1. Global Arnoldi process.
1. Set $V_1 = R^{(0)} / \left\| R^{(0)} \right\|$.
2. For $j = 1, 2, \ldots, k$ Do
3. $W := \mathcal{M}(V_j)$
4. For $i = 1, 2, \ldots, j$ Do
5. $h_{ij} := \langle W, V_i \rangle$
6. $W := W - h_{ij} V_i$
7. End for
8. $h_{j+1,j} := \|W\|$. If $h_{j+1,j} := 0$, then stop.
9. $V_{j+1} := W / h_{j+1,j}$
10. End for

Suppose that $\mathcal{V}_k = [V_1, V_2, \ldots, V_k]$ denotes the $m \times kpn$ matrix where

$$V_i = [V_1^{(i)}, V_2^{(i)}, \ldots, V_p^{(i)}],$$

for $i = 1, 2, \ldots, k$. Let $\overline{H}_k$ be a $(k+1) \times k$ upper Hessenberg matrix where its nonzero entries h_{ij} are computed by Algorithm 1 and H_k is the $k \times k$ matrix obtained from $\overline{H}_k$ by deleting its last row. It is not difficult to see that the matrix $\mathcal{V}_k$, produced by Algorithm 1, is an orthonormal basis for the $\mathcal{K}_k(\mathcal{M}, R^{(0)})$, i.e., $\mathcal{V}_k^T \circledcirc \mathcal{V}_k = I_k$.

The following proposition is easily deduced from Algorithm 1.

Proposition 3.1. *Let $\mathcal{V}_k$, $\overline{H}_k$ and H_k be defined as before, then we have the following relations:*

(1) $[\mathcal{M}(V_1), \mathcal{M}(V_2), \ldots, \mathcal{M}(V_k)] = \mathcal{V}_k((H_k \otimes I_p) \otimes I_n) + h_{k+1,k}[0_{m\times pn}, \ldots, 0_{m\times pn}, V_{k+1}]$.
(2) $[\mathcal{M}(V_1), \mathcal{M}(V_2), \ldots, \mathcal{M}(V_k)] = \mathcal{V}_{k+1}((\overline{H}_k \otimes I_p) \otimes I_n)$.

3.2. Gl-FOM for solving the general coupled linear matrix equations

Starting from an initial guess $X^{(0)} \in \mathbb{R}^{m\times pn}$ and the corresponding residual $R^{(0)} = C - \mathcal{M}(X^{(0)})$, the Gl-FOM algorithm computes the approximate solution $X^{(k)}$ such that

$$X^{(k)} \in X^{(0)} + \mathcal{K}_k(\mathcal{M}, R^{(0)}),$$

and

$$R^{(k)} = C - \mathcal{M}(X^{(k)}) \perp \mathcal{K}_k(\mathcal{M}, R^{(0)}). \tag{3.2}$$

Considering the orthonormal basis $\mathcal{V}_k = [V_1, V_2, \ldots, V_k]$ for $\mathcal{K}_k(\mathcal{M}, R^{(0)})$, we get

$$X^{(k)} = X^{(0)} + \sum_{i=1}^{k} V_i y_i^{(k)} = X^{(0)} + \mathcal{V}_k((y^{(k)} \otimes I_p) \otimes I_n), \tag{3.3}$$

where the real vector $y^{(k)} = [y_1^{(k)}, y_2^{(k)}, \ldots, y_k^{(k)}]^T$ is obtained by imposing the orthogonality condition (3.2).

Theorem 3.2. *The approximate solution $X^{(k)}$ produced by the Gl-FOM algorithm is given by $X^{(k)} = X^{(0)} + \mathcal{V}_k((y^{(k)} \otimes I_p) \otimes I_n)$ where $y^{(k)}$ is the solution of the following linear system*

$$H_k y = \beta e_1,$$

where $\beta = \left\| R^{(0)} \right\|$.

Proof. Straightforward computations show that

$$\begin{aligned}
R^{(k)} &= C - \mathcal{M}(X^{(k)}) \\
&= C - \mathcal{M}\left(X^{(0)} + \sum_{i=1}^{k} V_i y_i^{(k)} \right) \\
&= R^{(0)} - \sum_{i=1}^{k} \mathcal{M}(V_i) y_i^{(k)} \\
&= R^{(0)} - [\mathcal{M}(V_1), \ldots, \mathcal{M}(V_k)]\,((y^{(k)} \otimes I_p) \otimes I_n).
\end{aligned}$$

Using the first relation of Proposition 3.1, we derive

$$R^{(k)} = R^{(0)} - \left(\mathcal{V}_k((H_k \otimes I_p) \otimes I_n) + h_{k+1,k}[0_{m\times pn}, \ldots, 0_{m\times pn}, V_{k+1}]\right)((y^{(k)} \otimes I_p) \otimes I_n).$$

The orthogonality condition (3.2) implies that $\mathcal{V}_k^T \circledcirc R^{(k)} = 0$. Therefore, from the above relation and Eq. (2.2), we deduce

$$\mathcal{V}_k^T \circledcirc R^{(0)} = (\mathcal{V}_k^T \circledcirc \mathcal{V}_k)H_k y^{(k)}.$$

On the other hand, it is known that $\mathcal{V}_k^T \circledcirc \mathcal{V}_k = I_k$ and $R^{(0)} = \mathcal{V}_k((\beta e_1 \otimes I_p) \otimes I_n)$, and hence we can conclude the result immediately. □

The following proposition helps us to obtain the residual $\left\|R^{(k)}\right\|$ without computing $X^{(k)}$.

Proposition 3.3. *The norm of residual $R^{(k)}$ corresponding to the approximate solution $X^{(k)}$ computed by the Gl-FOM algorithm satisfies the following equality*

$$\left\|R^{(k)}\right\| = h_{k+1,k}\left|y_k^{(k)}\right|,$$

where $y_k^{(k)}$ is the last component of the vector $y^{(k)}$.

Proof. It is not difficult to see that

$$R^{(k)} = -h_{k+1,k}[0_{m\times pn}, \ldots, 0_{m\times pn}, V_{k+1}]((y^{(k)} \otimes I_p) \otimes I_n).$$

Now, the result can be easily derived by invoking the facts that $\left\|R^{(k)}\right\|^2 = (R^{(k)})^T \circledcirc R^{(k)}$ and $\|V_{k+1}\|^2 = V_{k+1}^T \circledcirc V_{k+1} = 1$. □

To save memory and CPU-time requirements, the Gl-FOM algorithm is used in a restarted mode. That is, the algorithm is restarted every k inner iterations, where k is a given fixed integer and the corresponding algorithm is denoted by Gl-FOM (k) and summarized as follows:

Algorithm 2. Gl-FOM(k) algorithm for Eq. (1.1).

1. Choose $X^{(0)}$ and a tolerance ε. Compute $R^{(0)} = C - \mathcal{M}(X^{(0)})$ and $V_1 = R^{(0)}$.
2. Construct the orthonormal basis $V_1, V_2, \ldots, V_k$ by Algorithm 1.
3. Find $y^{(k)}$ as the solution of the linear system

$$H_k y = \left\|R^{(0)}\right\| e_1.$$

4. Compute the residual $R^{(k)}$ and $\left\|R^{(k)}\right\|$ using Proposition 3.3.
5. If $\frac{\left\|R^{(k)}\right\|}{\left\|R^{(0)}\right\|} < \varepsilon$ Stop; else $R^{(0)} := R^{(k)}, V_1 := R^{(0)}$, go to 2.

3.3. Gl-GMRES for solving the general coupled linear matrix equations

Like the Gl-FOM algorithm the kth iterate $X^{(k)}$ of the Gl-GMRES algorithm belongs to the affine matrix Krylov subspace $X^{(0)} + \mathcal{K}_k(\mathcal{M}, R^{(0)})$. On the other hand, in the Gl-GMRES algorithm, the vector $y^{(k)}$ in Eq. (3.3) is obtained by imposing the following orthogonality condition

$$R^{(k)} = C - \mathcal{M}(X_k) \perp \mathcal{K}_k(\mathcal{M}, \mathcal{M}(R_0)). \tag{3.4}$$

The orthogonality condition (3.4) shows that $X^{(k)}$ can be obtained as the solution of the minimization problem

$$\min_{X - X^{(0)} \in \mathcal{K}_k(\mathcal{M}, R^{(0)})} \|C - \mathcal{M}(X)\|. \tag{3.5}$$

Now, we establish the following useful theorem.

Theorem 3.4. *The approximate solution $X^{(k)}$ computed by the Gl-GMRES algorithm is presented by $X^{(k)} = X^{(0)} + \mathcal{V}_k((y^{(k)} \otimes I_p) \otimes I_n)$ where $y^{(k)}$ is the solution of the following least square problem*

$$\min_{y\in\mathbb{R}^k} \left\|\beta e_1 - \overline{H}_k y\right\|_2, \tag{3.6}$$

where $\beta = \left\|R^{(0)}\right\|$.

Proof. Let $\mathcal{V}_k$ be the orthonormal basis for $\mathcal{K}_k(\mathcal{M}, R^{(0)})$ which is constructed by Algorithm 1. By some easy computations and using the second relation of Proposition 3.1, we have

$$\begin{aligned} R^{(k)} &= C - \mathcal{M}(X^{(k)}) \\ &= C - \mathcal{M}\left(X^{(0)} + \sum_{i=1}^{k} V_i y_i^{(k)}\right) \\ &= R^{(0)} - \sum_{i=1}^{k} \mathcal{M}(V_i) y_i^{(k)} \\ &= R^{(0)} - [\mathcal{M}(V_1), \ldots, \mathcal{M}(V_k)]\,((y^{(k)} \otimes I_p) \otimes I_n) \\ &= R^{(0)} - \mathcal{V}_{k+1}((\overline{H}_k \otimes I_p) \otimes I_n)((y^{(k)} \otimes I_p) \otimes I_n) \\ &= R^{(0)} - \mathcal{V}_{k+1}((\overline{H}_k y^{(k)} \otimes I_p) \otimes I_n). \end{aligned}$$

It is known that $R^{(0)} = \mathcal{V}_{k+1}((\beta e_1 \otimes I_p) \otimes I_n)$, hence

$$R^{(k)} = \mathcal{V}_{k+1}(((\beta e_1 - \overline{H}_k y^{(k)}) \otimes I_p) \otimes I_n).$$

Evidently $\left\|R^{(k)}\right\|^2 = (R^{(k)})^T \circledcirc R^{(k)}$, therefore using Eq. (2.2), we have

$$\begin{aligned} \left\|R^{(k)}\right\|^2 &= (\mathcal{V}_{k+1}(((\beta e_1 - \overline{H}_k y^{(k)}) \otimes I_p) \otimes I_n))^T \circledcirc (\mathcal{V}_{k+1}(((\beta e_1 - \overline{H}_k y^{(k)}) \otimes I_p) \otimes I_n)) \\ &= (\beta e_1 - \overline{H}_k y^{(k)})^T (\mathcal{V}_{k+1}^T \circledcirc \mathcal{V}_{k+1})(\beta e_1 - \overline{H}_k y^{(k)}), \end{aligned}$$

as $\mathcal{V}_{k+1}^T \circledcirc \mathcal{V}_{k+1} = I_{k+1}$, we get

$$\left\|R^{(k)}\right\|^2 = \left\|\beta e_1 - \overline{H}_k y^{(k)}\right\|_2^2.$$

Now, we can conclude the result from Eq. (3.5) immediately. $\square$

Consider the QR decomposition of the $(k+1) \times k$ matrix $\overline{H}_k$, i.e., $\overline{R}_k = Q_k \overline{H}_k$, where $\overline{R}_k$, Q_k are upper triangular and unity matrices, respectively. Assume that

$$\overline{g}_k = \left\|R^{(0)}\right\| Q_k e_1 = (\gamma_1, \gamma_2, \ldots, \gamma_{k+1})^T,$$

and R_k denotes the $k \times k$ matrix obtained from $\overline{R}_k$ by deleting its last row and g_k is the k-dimensional vector obtained from $\overline{g}_k$ by deleting its last component. Straightforward computations show that $y^{(k)} = R_k^{-1} g_k$.

The following theorem helps us to compute the norm of the kth residual in an inexpensive way.

Theorem 3.5. *The residual $R^{(k)} = C - \mathcal{M}(X^{(k)})$ obtained by the Gl-GMRES algorithm for the general coupled matrix equation satisfies the following equalities*

$$R^{(k)} = \gamma_{k+1} \mathcal{V}_{k+1}((Q_k^T e_{k+1} \otimes I_p) \otimes I_n),$$

and

$$\left\|R^{(k)}\right\| = |\gamma_{k+1}|,$$

where γ_{k+1} is the last component of the vector $\overline{g}_k$.

Proof. It is not difficult to see that

$$\begin{aligned} R^{(k)} &= R^{(0)} - \mathcal{V}_{k+1}((\overline{H}_k y^{(k)} \otimes I_p) \otimes I_n) \\ &= \mathcal{V}_{k+1}\left(\left(\left(\left\|R^{(0)}\right\| e_1 - \overline{H}_k y^{(k)}\right) \otimes I_p\right) \otimes I_n\right) \\ &= \mathcal{V}_{k+1}\left[\left(\left(Q_k^T Q_k \otimes I_p\right) \otimes I_n\right)\left(\left(\left(\left\|R^{(0)}\right\| e_1 - \overline{H}_k y^{(k)}\right) \otimes I_p\right) \otimes I_n\right)\right] \\ &= \mathcal{V}_{k+1}[((Q_k^T \otimes I_p) \otimes I_n)(((\overline{g}_k - \overline{R}_k y^{(k)}) \otimes I_p) \otimes I_n)]. \end{aligned}$$

As $y^{(k)} = R_k^{-1} g_k$, we get

$$R^{(k)} = \mathcal{V}_{k+1}[((Q_k^T \otimes I_p) \otimes I_n)((\gamma_{k+1} e_{k+1} \otimes I_p) \otimes I_n)] = \gamma_{k+1} \mathcal{V}_{k+1}((Q_k^T e_{k+1} \otimes I_p) \otimes I_n).$$

Evidently,

$$\left\|R^{(k)}\right\|^2 = (R^{(k)})^T \circledcirc R^{(k)} = \gamma_{k+1}^2 (Q_k^T e_{k+1})^T (\mathcal{V}_{k+1}^T \circledcirc \mathcal{V}_{k+1})(Q_k^T e_{k+1}) = \gamma_{k+1}^2,$$

which completes the proof. $\square$

Like the Gl-FOM algorithm, in application, the Gl-GMRES algorithm is restarted every k inner iterations, where k is a given fixed integer and the corresponding algorithm is denoted by Gl-GMRES (k) and presented as follows:

Algorithm 3. Gl-GMRES(k) algorithm for Eq. (1.1).

1. Choose $X^{(0)}$, a tolerance ε. Compute $R^{(0)} = C - \mathcal{M}(X^{(0)})$, and $V_1 = R^{(0)}$.
2. Construct the orthonormal basis $V_1, V_2, \ldots, V_k$ by Algorithm 1.
3. Determine $y^{(k)}$ as the solution of the least square problem:

$$\min_{y\in\mathbb{R}^k} \left\|\beta e_1 - \overline{H}_k y\right\|_2 .$$

   Compute $X^{(k)} = X^{(0)} + \mathcal{V}_k((y^{(k)} \otimes I_p) \otimes I_n)$.

4. Compute the residual $R^{(k)}$ and $\left\|R^{(k)}\right\|$ using Theorem 3.5.
5. If $\frac{\left\|R^{(k)}\right\|}{\left\|R^{(0)}\right\|} < \varepsilon$ Stop; else $R^{(0)} := R^{(k)}$, $V_1 := R^{(0)}$, go to 2.

As we have seen, the norm of the residual obtained by the Gl-FOM is not minimized at each step and hence it may oscillate, but the Gl-GMRES algorithm overcomes this drawback. Similar to the classical GMRES and FOM algorithms, if $\mathcal{M}$ is symmetric then the Gl-GMRES and Gl-FOM algorithms result in the Global Conjugate Residual and Global Conjugate Gradient algorithms, respectively.

4. Numerical examples

In this section, some numerical examples are presented to illustrate the effectiveness of the Gl-GMRES(k) and Gl-FOM(k) algorithms to solve (1.1). All the numerical experiments presented in this section were computed in double precision with some Matlab codes on a Pentium 4 PC, with a 3.06 GHz CPU and 1.00GB of RAM.

Example 4.1. For this experiment, we consider the general coupled matrix equations

$$\begin{cases} AX_1 + X_2B = C_1, \\ BX_1 + X_2A = C_2, \end{cases}$$

where

$$A = \begin{pmatrix} 4 & -1 & & -1 \\ -1 & 4 & \ddots & \\ & \ddots & \ddots & -1 \\ -1 & & -1 & 4 \end{pmatrix}, \quad \text{and} \quad B = \begin{pmatrix} 8 & -2 & & -2 \\ -2 & 8 & \ddots & \\ & \ddots & \ddots & -2 \\ -2 & & -2 & 8 \end{pmatrix},$$

are $m \times m$ matrices. The right-hand side of the corresponding system $\mathcal{M}(X) = C$ was taken such that $X = (X_1, X_2)$ is the exact solution of the system where $X_1 = \text{tridiag}(1, 1, 1)$ and $X_2 = \text{tridiag}(1, -1, 1)$. The initial guess was taken to be zero and the test was stopped as soon as

$$\frac{\left\|R^{(j)}\right\|}{\left\|R^{(0)}\right\|} < 10^{-8},$$

where $R^{(j)} = C - \mathcal{M}(X^{(j)})$. The numerical results are given in Table 1. In this table, "iters" and "Err" stand for the number of iterations needed for the convergence and

$$\text{Err} = \|(X_1, X_2) - (\bar{X}_1, \bar{X}_2)\|_\infty,$$

respectively, where $(\bar{X}_1, \bar{X}_2)$ is the approximate solution computed by Algorithm 3. As we observe, numerical results show that the Gl-GMRES(5) and Gl-FOM(5) algorithms are efficient for solving the general coupled matrix equations and the results of the Gl-GMRES(5) algorithm are better than those of the Gl-FOM(5) algorithm.

Example 4.2. Let

$$T_{d,k} = \text{tridiag}\left(-1 + \frac{10}{k+1}, d, -1 + \frac{10}{k+1}\right) \in \mathbb{R}^{k\times k}.$$

We consider the general coupled matrix equations

$$\begin{cases} A_{11}X_1B_{11} + A_{12}X_2B_{12} = C_1, \\ A_{21}X_1B_{21} + A_{22}X_2B_{22} = C_2, \end{cases}$$

Table 1
Numerical results for Example 4.1.

m	Gl-GMRES(5)		Gl-FOM(5)	
	Iters	Err	Iters	Err
250	21	2.02e−6	23	2.96e−6
500	20	5.28e−6	23	4.47e−6
750	20	5.86e−6	23	6.30e−6
1000	20	6.32e−6	23	7.97e−6

Table 2
Numerical results for Example 4.2.

n	Gl-GMRES(5)		Gl-FOM(5)	
	Iters	Err	Iters	Err
300	82	1.37e−6	145	2.46e−6
600	86	3.92e−6	160	3.12e−6
900	87	4.90e−6	168	4.28e−6

where $B_{11} = B_{22} = T_{2,n}, B_{12} = B_{21} = T_{3,n}$ and $A_{11} = A_{12} = A_{21} = A_{22} =$ GR3030, in which GR3030 has been downloaded from the Matrix-Market website [19]. Here we mention that GR3030 is a matrix of order 900 with 4322 nonzero entries. The right-hand side of the corresponding system $\mathcal{M}(X) = C$ was taken such that $X = (X_1, X_2)$ is the exact solution of the system where

$$(X_1)_{ij} = \begin{cases} 1, & |i-j| \le 1, \\ 0, & \text{otherwise}, \end{cases}$$

$$(X_2)_{ij} = \begin{cases} 1, & i = j, \\ 0, & \text{otherwise}. \end{cases}$$

All of other assumptions are as the previous example. The numerical results for different values of n are presented in Table 2. As seen, numerical results demonstrate that the Gl-GMRES(5) and Gl-FOM(5) algorithms are profitable for solving the general coupled matrix equations. Another observation which can be posed here is that the Gl-GMRES(5) algorithm works better than the Gl-FOM(5) algorithm.

Example 4.3. In this example, we compare the numerical results of the Gl-GMRES(k) and Gl-FOM(k) algorithms with those of the method presented in [14]. To do so, we choose the second example in [14] where the following coupled linear matrix equations

$$\begin{cases} AX + YB = C, \\ DX + YE = F, \end{cases} \tag{4.1}$$

are considered, in which

$$A = \begin{pmatrix} 2.00 & 1.00 \\ -1.00 & 2.00 \end{pmatrix}, \quad B = \begin{pmatrix} 1.00 & -0.20 \\ 0.20 & 1.00 \end{pmatrix}, \quad D = \begin{pmatrix} -2.00 & -0.50 \\ 0.50 & 2.00 \end{pmatrix},$$
$$E = \begin{pmatrix} -1.00 & -3.00 \\ 2.00 & -4.00 \end{pmatrix}, \quad C = \begin{pmatrix} 13.2 & 10.60 \\ 0.60 & 8.40 \end{pmatrix}, \quad F = \begin{pmatrix} -9.50 & -18.00 \\ 16.00 & 3.50 \end{pmatrix}.$$

The unique exact solution of system (4.1) is given by (X^*, Y^*), where

$$X^* = \begin{pmatrix} 4.00 & 3.00 \\ 3.00 & 4.00 \end{pmatrix}, \quad Y^* = \begin{pmatrix} 2.00 & 1.00 \\ -2.00 & 3.00 \end{pmatrix}.$$

In this example, we use $X^{(0)} = Y^{(0)} = 10^{-6} I_{2\times 2}$ as the initial guess and

$$\delta_k = \frac{\|X^{(k)} - X^*\|_F + \|Y^{(k)} - Y^*\|_F}{\|X^*\|_F + \|Y^*\|_F} < 10^{-8}$$

as the stopping criterion. In Table 3, the numerical results of the Gl-GMRES(4) and Gl-FOM(4) algorithms together with the gradient based iterative method presented in [14] were given. As before, "iters" stands for the number of iterations for the convergence. In this table, the CPU time (in seconds) for computing the approximate solution were also given. As observed, for this example the numerical results in terms of both number of iterations and CPU-time(s) for the Gl-GMRES(4) and Gl-FOM(4) algorithms are better than those of the gradient based iterative method proposed in [14]. We believe that the Gl-GMRES(k) and Gl-FOM(k) algorithms in general are more suitable than the gradient based algorithm given in [14], especially for large problems.

Table 3
Numerical results for Example 4.3.

	Gl-GMRES(4)	Gl-FOM(4)	Gradient based [14]
Iters	28	54	226
CPU time	0.047	0.078	0.125

5. Conclusion

We have extended the global FOM and GMRES algorithms to solve the general coupled matrix equations. Furthermore, by introducing a new matrix product, the global FOM and GMRES algorithms have been analyzed for solving the general coupled matrix equations. Moreover, some numerical results of the global GMRES and FOM algorithms have been presented. Our numerical experiments have illustrated the effectiveness of these algorithms for solving general coupled matrix equations. More theoretical results of the proposed algorithms are under investigation.

Acknowledgments

The authors are grateful to the anonymous referees for their comments which substantially improved the quality of this paper.

References

[1] A. Bouhamidi, K. Jbilou, A note on the numerical approximate solutions for generalized Sylvester matrix equations with applications, Appl. Math. Comput. 206 (2008) 687–694.
[2] K. Jbilou, A. Messaudi, H. Sadok, Global FOM and GMRES algorithms for matrix equations, Appl. Numer. Math. 31 (1999) 49–63.
[3] K. Jbilou, A.J. Riquet, Projection methods for large Lyapunov matrix equations, Linear Algebra Appl. 415 (2006) 344–358.
[4] J.J. Zhang, A note on the iterative solutions of general coupled matrix equation, Appl. Math. Comput. 217 (2011) 8386–9380.
[5] B. Zhou, G.R. Duan, On the generalized Sylvester mapping and matrix equation, Systems Control Lett. 57 (3) (2008) 200–208.
[6] Z.Y. Li, Y. Wang, Iterative algorithm for minimal norm least squares solution to general linear matrix equations, Int. J. Comput. Math. 87 (2010) 2552–2567.
[7] X.W. Chang, J.S. Wang, The symmetric solution of the matrix equations $AX + YA = C$, $AXA^T + BYB^T = C$ and $(A^TXA, B^TXB) = (C, D)$, Linear Algebra Appl. 179 (1993) 171–189.
[8] Q.W. Wang, J.H. Sun, S.Z. Li, Consistency for bi(skew)symmetric solutions to systems of generalized Sylvester equations over a finite central algebra, Linear Algebra Appl. 353 (2002) 169–182.
[9] M. Dehghan, M. Hajarian, The general coupled matrix equations over generalized bisymmetric matrices, Linear Algebra Appl. 432 (2010) 1531–1552.
[10] Y. Saad, M.H. Schultz, GMRES: a generalized minimal residual method for solving nonsymmetric linear systems, SIAM J. Sci. Statist. Comput. 7 (1986) 856–869.
[11] F. Ding, P.X. Liu, J. Ding, Iterative solutions of the generalized Sylvester matrix equations by using the hierarchical identification principle, Appl. Math. Comput. 197 (2008) 41–50.
[12] F. Ding, T. Chen, On iterative solutions of general coupled matrix equations, SIAM J. Control Optim. 44 (2006) 2269–2284.
[13] B. Zhou, Z.Y. Li, G.R. Duan, Y. Wang, Weighted least squares solutions to general coupled Sylvester matrix equations, J. Comput. Appl. Math. 224 (2009) 759–776.
[14] B. Zhou, G.R. Duan, Z.Y. Li, Gradient based iterative algorithm for solving coupled matrix equations, Systems Control Lett. 58 (2009) 327–333.
[15] Y. Saad, Iterative Methods for Sparse Linear Systems, PWS press, New York, 1995.
[16] D.K. Salkuyeh, F. Toutounian, New approaches for solving large Sylvester equations, Appl. Math. Comput. 173 (2006) 9–18.
[17] Y.Q. Lin, Implicitly restarted global FOM and GMRES for nonsymmetric equations and Sylvester equations, Appl. Math. Comput. 167 (2005) 1004–1025.
[18] R. Bouyouli, K. Jbilou, R. Sadaka, H. Sadok, Convergence properties of some block Krylov subspace methods, J. Comput. Appl. Math. 196 (2006) 498–511.
[19] Matrix Market, http://math.nist.gov/MatrixMarket, August 2005.